CHRS PERSIANN-CONNECT Dataset

Global satellite and remote sensing observations capture an enormous amount of data for essentially every precipitation event that occurs. Using PERSIANN precipitation estimates, we define individual precipitation events as segmented or "connected" objects in four-dimensional (4D) space (latitude, longitude, time and intensity (mm/hr)). This dataset is called PERSIANN-CONNECT. We store the segmented precipitation objects in a publicly available ftp directory and PostgreSQL database. The database can be used to subset and download precipitation objects.

The global version of this dataset can be accessed at http://connect.eng.uci.edu

Applications of the data:

- Precipitation event visualization and comparison
  - Search for specific events or characteristics
- Extreme precipitation event analysis
  - Create a data subset to include extreme events for a particular region
- Weather and climate studies
  - Gain insight into the changes in precipitation event characteristics
- Engineering Applications
  - Use Objects for "storm design" or precipitation scenario development in hydrological modeling and water resource planning.

Please see the below article for more information:

Sellars, S., P. Nguyen, W. Chu, X. Gao, K. Hsu and S. Sorooshian (2013), Computational Earth Science: Big Data Transformed Into Insight, Eos Trans. AGU, 94(32), 277.

Sellars, S., X. Gao and S. Sorooshian (2015), An Object-Oriented Approach to Investigate Impacts of Climate Oscillations on Precipitation: A Western United States Case Study. J. Hydrometeor, 16, 830–842.

Nguyen, P., S. Sellars, A. Thorstensen, Y. Tao, H. Ashouri, D. Braithwaite, K. Hsu and S. Sorooshian (2014), Satellites Track Precipitation of Super Typhoon Haiyan, Eos Trans. AGU, 95(16), 133.

What is a precipitation object?

When you define an item as an object, you can then describe it using carefully chosen attributes or features. The objects, and the attributes found in the empirical data, can then be analyzed using statistical methods. These connected precipitation objects also allow for the investigation and analysis of the evolution of precipitation events in space and time as estimated from satellite observations. A user of this object-oriented method, depending on their particular area of interest, can define their own attributes. In our case, we have chosen simple physical characteristics typically used to describe earth science events. The precipitation objects, as we have defined them, are considered experimental and part of a data intensive research project. Please be patient as aspects of this dataset will change and evolve.

What is our Purpose?

Many advances in the theoretical understanding of atmospheric and oceanic dynamics have been based on numerical modeling. We seek additional advances by harnessing the vast amounts of remote sensing

information and focusing on a data driven and computer vision approach, which looks at observational and modeled data from a 4D perspective. By viewing precipitation events as objects in space and time, we provide enhanced information regarding the evolution of the event itself, as well as its characteristics.

The Connectivity Algorithm

The connectivity algorithm is designed to ensure that all pixels of precipitation estimates are connected (assuming a set intensity threshold) in both space and time, allowing for the feature to be analyzed as a 4D object. Using Matlab, a code was developed by researchers at CHRS for precipitation event segmentation. A precipitation voxel is defined as a volumetric unit of precipitation. The algorithm uses 4D space (longitude, latitude, time and intensity) and looks for connected voxels in space and time. If two voxels have at least one connection (face, corner or edge) in space or time, they can be grouped into a common object. The algorithm segments all of the objects, labels them, and puts them in a descriptive table. We exclude a large number of small precipitation events by setting a few algorithm criteria. In this version of the algorithm, we apply a threshold of precipitation intensity at 1 millimeter per hour (mm/hr.), precipitation duration at 24 hours, and limit the number of connections in space and time to 6 (i.e. faces of the voxel). Future versions of this dataset will experiment with different combinations of these criteria to determine the optimal settings for viewing precipitation as objects.

Data

The data we used for this research dataset is CHRS's PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (See Hsu et al., [1997]; Sorooshian et al., [2000])), bias corrected, hourly, .25-degree precipitation dataset from March 1st, 2000 to January 1st, 2011. This bias adjusted product maintains total monthly precipitation estimates, consistent with GPCP (Global Precipitation Climatology Project) product. The dataset retains the spatial and temporal features of precipitation estimates provided by the original PERSIANN algorithm at 0.25-degree spatial and 1-hourly temporal resolution.

Raw Gridded Format:
4-byte binary float (little-endian byte order). Units: mm/hr
Spatial coverage is: 60N to -60S latitude and 0 to 360 longitude
Spatial Resolution: .25 x .25 degree
Geometry: 480 rows x 1440 columns
The data is stored in C style row centric format. With the first value centered at 59.875, .125, the second value at 59.875, .375, and the last 2 values are centered at: -59.875, 359.625 and -59.875, 359.875

The three hourly version of this data can be accessed here:
http://fire.eng.uci.edu/PERSIANN/adj_persiann_3hr.html.

4D Object Format:
Data is stored in comma separate value format.
Each row contains the voxel information: Object ID, Latitude (degree), Longitude (degree), Time (date) and Intensity (mm/hour).
Spatial coverage is same as Raw Gridded Data: 60N to -60S latitude 0 to 360 longitude
Voxel Resolution: .25 x .25 degree

Segmented 4D Object Characteristic Data:
Nxd Matrix format

N is the # of objects in rows, d is the # of characteristics in columns
Current descriptive statistics:
1) Object ID
2) Average Intensity (mm/hr)
3) Median Intensity (mm/hr)
4) Maximum Intensity (mm/hr)
5) Standard Deviation
6) Volume (m^3)
7) Duration
8) Centroid ("center of mass") Latitude
9) Centroid ("center of mass") Longitude
10) Velocity or Speed (km/hr) - Calculated using the distance from the starting centroid to the ending centroid.
11) Starting Time
12) Ending Time
13) Starting ("center of mass") Centroid Latitude
14) Starting ("center of mass") Centroid Longitude
15) Ending ("center of mass") Centroid Latitude
16) Ending ("center of mass") Centroid Longitude

Climate Index Data:

Monthly climate indices including AO, MJO, ENSO, SST from NOAA are also incorporated with each precipitation event in the statistics table.

Words of Caution

These objects, as we have defined them, contain large amounts of information, dynamics and physics that are fundamental to describing precipitation objects, but are not fully represented in the data structure we have used. Some objects contain hundreds to thousands of smaller precipitation events. At this time, the criteria used to segment precipitation objects can be thought of as exploratory. The criteria used to produce PERSIANN-CONNECT always has room for improvement. For example, the minimum threshold criteria (e.g. 1mm/hr for each voxel and 24 hour of existence) is the current settings used to define precipitation objects. Future data releases may use other thresholds, which prove to be more effective at describing precipitation events. These types of criteria for defining precipitation objects were investigated by Takeuchi [1985] (among others). He looked at precipitation objects for short-lived heavy rainfall areas, concluding that spatial resolution and temporal resolution (i.e. 15mins vs 1hr) alter the mean characteristic of the results. Please keep this in mind during your exploration of the objects.

Satellite based precipitation estimates are extremely useful for monitoring precipitation over regions with no ground based observational capability. It should be noted that these precipitation estimates have errors and uncertainty. Satellite based precipitation estimation algorithms continue to be developed and improved, and we are very excited for the launch of NASA's GPM mission that will provide more accurate information. If you are interested in the accuracy of satellite based precipitation products, please view these verification sites below.

Support

- California Department of Water Resources

References

Sellars, S., P. Nguyen, W. Chu, X. Gao, K. Hsu and S. Sorooshian (2013), Computational Earth Science: Big Data Transformed Into Insight, Eos Trans. AGU, 94(32), 277.

Hsu, K., X. Gao, and S. Sorooshian (1997), Precipitation estimation from remotely sensed information using artificial neural networks, J. Appl. Meteorol., 36, 1176-1190, doi:10.1175/ 1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.

Sorooshian, S., K. Hsu, X. Gao, H. Gupta, B. Imam, and D. Braithwaite (2000), Evaluation of PERSIANN system satellite-based estimates of tropical rainfall, Bull. Am. Meteorol. Soc., 81, 2035-2046, doi:10.1175/1520-0477(2000)081<2035 :EOPSSE>2.3.CO;2.

Takeuchi, K. (1985). An Automatic Storm Tracking Method Used To Analyze Traveling Characteristics Of Heavy Rain Areas. Natural disaster science, 7(1), 13-24.